Contemporary Mathematics
http://ojs.wiserpub.com/index.php/CM/

Research Article

Results on Enhanced Continuous Random Variable's Probability Distribution Using a Different Exponential Function for the Normal Probability Distribution

C V Rao

Department of General Studies, Jubail Industrial College, P.O. Box-10099, Jubail Industrial City, 31961, Saudi Arabia
E-mail: cvrao@rcjy.edu.sa

Received: 28 August 2024; **Revised:** 14 October 2024; **Accepted:** 15 November 2024

Abstract: In the history of probability and statistics, general normal probability has played an important role. I utilize a different exponential function to create a new continuous probability density function akin to the Laplace Gauss distribution function. Finding an alternate probability distribution is required so that the probability can be determined without referring to the table data. I noticed that the findings are pretty comparable to the normal probability distribution values. I checked the new results against a few standard cases. The advantage of this exponential function is that we can calculate the probability of a z value with more than two decimals, whereas with a normal distribution, we can only use z values with two decimals.

Keywords: random variable, normal probability distribution, error function, hyperbolic secant function, cumulative probability distribution

MSC: 00A35, 11k06, 11k41, 11k99

1. Introduction

The theory of probability distribution is crucial in the fields of statistics and probability. Relevant probability distributions in modern mathematics include the exponential, gamma, Cauchy, and normal (Laplace-Gaussian) distributions. Finally, there is the χ^2 distribution. The general normal probability distribution [1], [2, Chapter 26, equation 26.2.12], [3–20] is the most practical distribution.

1.1 *General normal distribution*

One important and frequently used probability distribution in statistics is the normal distribution, also referred to as the Gaussian distribution. It describes the distribution of values in a random variable. The majority of data points in a dataset with a normal distribution are clustered around the mean, or average, value, and the likelihood of finding values further from the mean decreases uniformly on both sides.

Key Features of the Normal Distribution

DOI: https://doi.org/10.37256/cm.6120255591

1. The normal distribution has a symmetric, bell-shaped curve that centers on the mean. This curve is known as the probability density function (PDF).

2. In a normal distribution, the mean, median, and mode of data all align at the center.

3. Symmetry: The normal distribution is symmetric around the mean. This signifies that the distribution's left and right sides are symmetrical.

4. Asymptotic: The tails of the normal distribution curve approach but do not touch the horizontal axis. This means that there is always some chance, however tiny, of seeing values that deviate from the mean.

5. Defined by two parameters:

(i) The mean (μ) represents the center of the distribution. It decides where the curve's peak is.

(ii) The standard deviation (σ) is a measure of the distribution's spread. A lower standard deviation indicates that the data points are closer to the mean, resulting in a steeper curve, whereas a higher standard deviation spreads the data points out, flattening the curve.

Application of Normal Distribution:

The Central Limit Theorem (CLT), which asserts that regardless of the starting distribution of the variables, the sum (or average) of a large number of independent, identically distributed random variables tends to be normal, places special emphasis on the normal distribution. Even in cases when the data did not originally follow a normal distribution, this theorem validates the use of the normal distribution in many statistical procedures.

Probability Density Function (PDF)

Standard normal density function

$$n(x;\ \mu,\ \sigma) = \frac{1}{\sigma\sqrt{2\pi}} exp\left(\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)$$

where:

- x is the random variable.
- μ is the mean.
- σ is the standard deviation.
- exp denotes the exponential function.

Cumulative Distributive Function:

Standard normal density function

$$\emptyset(z) = \frac{1}{\sqrt{2\pi}} e^{\frac{-1}{2}z^2},$$

where $z = \dfrac{x-\mu}{\sigma}$.

The standard normal cumulative distribution function (c.d.f) $\Phi(z)$ gives the area to the left of z under the standard normal curve

$$Phi(z)\ = \int_{-\infty}^{z} \emptyset(y)\, dy$$

total area $\dfrac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\frac{-1}{2}(z)^2} dz = 1.$

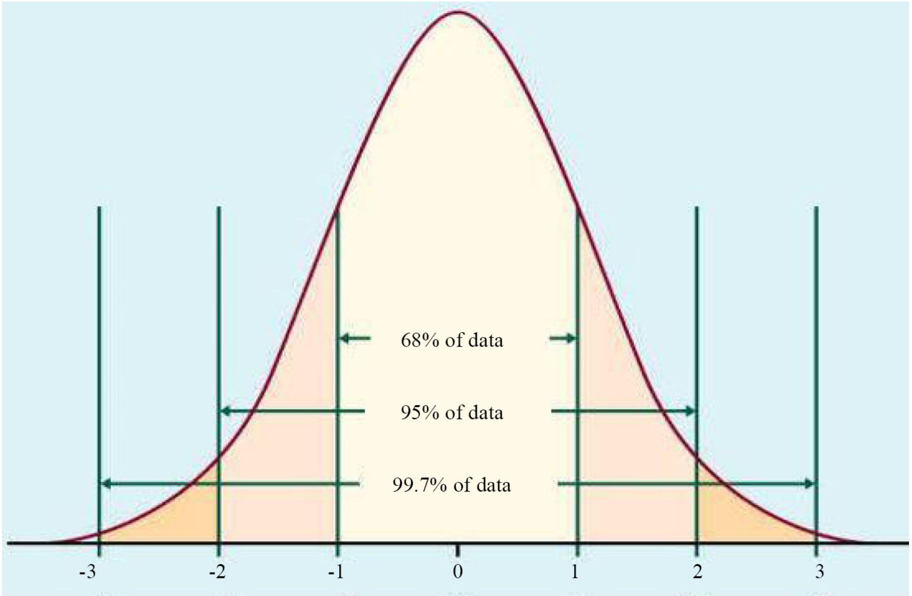

Figure 1. Normal distribution curve

Figure 1 shows the probability of being within one standard deviation of the mean.

$$\Phi(-1,\ 1) \approx 68.3\%$$

The likelihood is within two standard deviations of the mean.

$$\Phi(-2,\ 2) \approx 95.4\%$$

Probability is within three standard deviations of the mean.

$$\Phi(-3,\ 3) \approx 99.7\%$$

By the symmetry of the normal curve

$$\Phi(-z) = 1 - \Phi(z)\,(-\infty < z < \infty)$$

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{0} e^{\frac{-1}{2}(z)^2} dz = 0.5$$

also we know that.

The probability of the interval $(a,\ b)$ in the standard normal distribution is

$$\Phi(a,\ b) = \Phi(b) - \Phi(a)$$

By the difference rule of probabilities from Figure 1

$$\Phi(-z,\ z) = \Phi(z) - \Phi(-z) = \Phi(z) - (1 - \Phi(z)) = 2\Phi(z) - 1$$

Theorem 1 The mean and variance of $n(x,\ \mu,\ \sigma)$ are μ and σ^2.
Proof. To evaluate the mean, we first calculate

$$E(X-\mu) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (x-\mu)\, e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

Let $t = \dfrac{x-\mu}{\sigma} \Rightarrow dt = \dfrac{dx}{\sigma}$

$$E(X-\mu) = \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} t e^{\frac{-1}{2}(t)^2} dt$$

The function $f(t) = te^{\frac{-1}{2}(t)^2}$ is an odd one.
So $\int_{-\infty}^{\infty} te^{\frac{-1}{2}(t)^2} dt = 0$
Therefore $E(X-\mu) = 0$

$$E(X) = \mu$$

The normal distribution's variance is

$$E\left[(X-\mu)^2\right] = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (x-\mu)^2 e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

Let $t = \dfrac{x-\mu}{\sigma}$

$$\Rightarrow dt = \frac{dx}{\sigma}$$

$$E(X-\mu) = \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} t^2 e^{\frac{-1}{2}(t)^2} dt$$

Here $g(t) = t^2 e^{\frac{-1}{2}(t)^2}$ is an even function

$$E\left[(X-\mu)^2\right] = 2\frac{\sigma^2}{\sqrt{2\pi}} \int_0^\infty t^2 e^{\frac{-1}{2}(t)^2} dt$$

By using integration by parts we get

$$E\left[(X-\mu)^2\right] = \sigma^2\sqrt{\frac{\pi}{2}}\left\{\left[-te^{\frac{-1}{2}(t)^2}\right]_0^\infty + \int_0^\infty e^{\frac{-1}{2}(t)^2} dt\right\}$$

$$E\left[(X-\mu)^2\right] = \sigma^2\sqrt{\frac{\pi}{2}}\left\{0+\sigma^2\sqrt{\frac{\pi}{2}}\right\}$$

$$E\left[(X-\mu)^2\right] = \sigma^2$$

□

2. Improved probability function

We can choose Probability density function as

$$n(x,\ \mu,\ \sigma) = \psi(x) = \frac{\sqrt{3}e^{\sqrt{3}\left(\frac{x-\mu}{\sigma}\right)}}{\sigma\left(1+e^{\sqrt{3}\left(\frac{x-\mu}{\sigma}\right)}\right)^2}$$

Where μ is the mean and σ is the Standard deviation

$$\psi(z) = \frac{\sqrt{3}e^{\sqrt{3}z}}{\left(1+e^{\sqrt{3}z}\right)^2}$$

Cumulative Distribution function

$$\Psi(z) = \int_{-\infty}^{z} \psi(y)\, dy$$

Total area $\int_{-\infty}^{\infty} \frac{\sqrt{3}e^{\sqrt{3}z}}{\left(1+e^{\sqrt{3}z}\right)^2} dz = 1$

Improved form of cumulative distributive function

$$\Psi(z) = \sqrt{3}\int_{-\infty}^{z} \frac{e^{\sqrt{3}y}}{\left(1+e^{\sqrt{3}y}\right)^2} dy$$

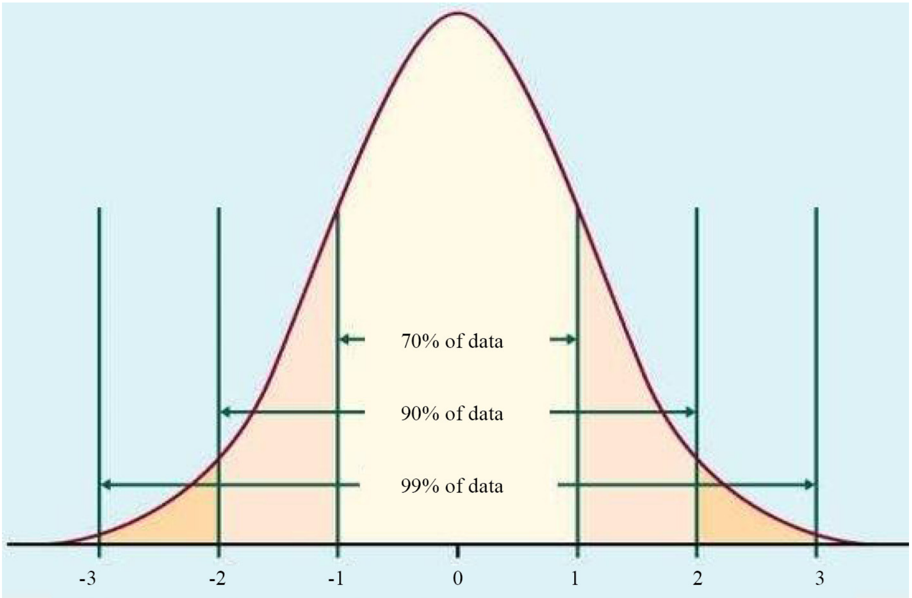

Figure 2. Improved form of probability distribution curve

Figure 2 shows the probability of being within one standard deviation of the mean.

$$\Psi(z)\ (-1,\ 1) = 69.93\%$$

Probability is within two standard deviations of the mean.

$$\Psi(z)\ (-2,\ 2) = 90\ \%$$

Probability is within three standard deviations of the mean.

$$\Psi(z)\ (-3,\ 3) = 98.89\%$$

Note :

$$1.\ P(-\infty < Z < \infty) = 1$$

$$2.\ P(Z < z) = \frac{e^{\sqrt{3}z}}{\left(1 + e^{\sqrt{3}z}\right)}$$

3. $P(Z > z) = 1 - \frac{e^{\sqrt{3}z}}{\left(1 + e^{\sqrt{3}z}\right)} = \frac{1}{\left(1 + e^{\sqrt{3}z}\right)}$

4. $P(z_1 < Z < z_2) = \frac{e^{\sqrt{3}z_2}}{\left(1 + e^{\sqrt{3}z_2}\right)} - \frac{e^{\sqrt{3}z_1}}{\left(1 + e^{\sqrt{3}z_1}\right)} = \frac{1}{\left(1 + e^{\sqrt{3}z_1}\right)} - \frac{1}{\left(1 + e^{\sqrt{3}z_2}\right)}$

Theorem 2 The mean and variance of $c(x;\ \mu,\ \sigma)$ are μ and σ^2.
Proof. To find the mean, we first calculate

$$E(X - \mu) = \frac{\sqrt{3}}{\sigma} \int_{-\infty}^{\infty} \frac{(x - \mu)\, e^{\sqrt{3}\left(\frac{x-\mu}{\sigma}\right)}}{\left(1 + e^{\sqrt{3}\left(\frac{x-\mu}{\sigma}\right)}\right)^2} dt$$

Let $t = \frac{x - \mu}{\sigma} \Rightarrow dt = \frac{dx}{\sigma}$

$$E(X - \mu) = \frac{\sigma^2 \sqrt{3}}{\sigma} \int_{-\infty}^{\infty} \frac{t e^{\sqrt{3}t}}{\left(1 + e^{\sqrt{3}t}\right)^2} dt$$

$$u(t) = \frac{t e^{\sqrt{3}t}}{\left(1 + e^{\sqrt{3}t}\right)^2}$$

$$u(-t) = \frac{-t e^{-\sqrt{3}t}}{\left(1 + e^{-\sqrt{3}t}\right)^2} = \frac{-t e^{-\sqrt{3}t}}{e^{-2\sqrt{3}t}\ \left(1 + e^{\sqrt{3}t}\right)^2} = \frac{-t e^{\sqrt{3}t}}{\left(1 + e^{\sqrt{3}t}\right)^2} = -u(t)$$

$u(t)$ is an odd function

$$\int_{-\infty}^{\infty} \frac{t e^{\sqrt{3}t}}{\left(1 + e^{\sqrt{3}t}\right)^2} dt = 0$$

Therefore $E(X - \mu) = 0$.
The variance of the normal distribution is given by

$$E\left[(X - \mu)^2\right] = \frac{\sqrt{3}}{\sigma} \int_{-\infty}^{\infty} \frac{(x - \mu)^2 e^{\sqrt{3}\left(\frac{x-\mu}{\sigma}\right)}}{\left(1 + e^{\sqrt{3}\left(\frac{x-\mu}{\sigma}\right)}\right)^2} dx$$

Let $t = \frac{x - \mu}{\sigma} \Rightarrow dt = \frac{dx}{\sigma}$

$$E\left[(X-\mu)^2\right] = \frac{\sqrt{3}}{\sigma}\int_{-\infty}^{\infty}\frac{(\sigma t)^2 e^{\sqrt{3}t}}{\left(1+e^{\sqrt{3}t}\right)^2}\sigma dt$$

$$E\left[(X-\mu)^2\right] = \sqrt{3}\sigma^2\int_{-\infty}^{\infty}\frac{t^2 e^{\sqrt{3}t}}{\left(1+e^{\sqrt{3}t}\right)^2}dt$$

$$v(t) = \frac{t^2 e^{\sqrt{3}t}}{\left(1+e^{\sqrt{3}t}\right)^2}$$

$$v(-t) = \frac{t^2 e^{-\sqrt{3}t}}{\left(1+e^{-\sqrt{3}t}\right)^2} = \frac{t^2 e^{-\sqrt{3}t}}{e^{-2\sqrt{3}t}\left(1+e^{\sqrt{3}t}\right)^2} = \frac{t^2 e^{\sqrt{3}t}}{\left(1+e^{\sqrt{3}t}\right)^2} = v(t)$$

$v(t)$ is an even function

$$\Rightarrow E\left[(X-\mu)^2\right] = 2\sqrt{3}\sigma^2\int_0^{\infty}\frac{t^2 e^{\sqrt{3}t}}{\left(1+e^{\sqrt{3}t}\right)^2}dt$$

By using integration by parts we get

$$E\left[(X-\mu)^2\right] = 2\sqrt{3}\sigma^2\left\{\left[\frac{-t^2}{\sqrt{3}\left(1+e^{\sqrt{3}t}\right)}\right]_0^{\infty} + \int_0^{\infty}\frac{2t}{\sqrt{3}\left(1+e^{\sqrt{3}t}\right)}dt\right\}$$

$$E\left[(X-\mu)^2\right] = 2\sqrt{3}\sigma^2\left\{0+\frac{2}{\sqrt{3}}\int_0^{\infty}\frac{t}{\left(1+e^{\sqrt{3}t}\right)}dt\right\}$$

$$E\left[(X-\mu)^2\right] = 4\sigma^2\int_0^{\infty}\frac{t}{\left(1+e^{\sqrt{3}t}\right)}dx$$

$$E\left[(X-\mu)^2\right] = \sigma^2\frac{\pi^2}{9} = 1.09662\,\sigma^2 \approx \sigma^2$$

Variance = σ^2. □

Problem 1 The GRE score is created by $N(48, 194)$. The required score for admission to the specified college is 205.

(a) What is the percentage of students admitted? If 25,000 students took the GRE?

(b) How many of them received a score higher than 200?

Sol:

(a) Given $N\left(\mu,\ \sigma^2\right) = N(48,\ 194)$ mean $\mu = 194$ and standard deviation $\sigma = 48$ Probability of the student's GRE score 205:

$$z_1 = \frac{x_1 - \mu}{\sigma} = \frac{205 - 194}{48} = 0.2291$$

$$\mathrm{P}(Z > z_1) = \frac{1}{\left(1 + e^{\sqrt{3}z_1}\right)} = 0.402$$

Therefore, percentage of students are eligible to join = 40.2%.

From Normal distribution method

$$P(Z > z_1) = P(Z > 0.23) = 1 - 0.5910 = 0.4090$$

Therefore, percentage of students are eligible to join $= 40.9\%$.

(b) GProbability of students having more than 200:

$$z_2 = \frac{x_2 - \mu}{\sigma} = \frac{200 - 194}{48} = 0.125$$

$$P(Z > z_2) = \frac{1}{\left(1 + e^{\sqrt{3}z_2}\right)} \approx 0.45$$

Number of expected gifted students

$$np = 25,000 \times 0.45 = 11,250$$

From Normal distribution method $P(Z > 0.23) = 1 - 0.5478 \approx 0.45$

$$np = 25,000 \times 0.45 = 11,250$$

Problem 2 The annual average number of accidents in Bombay is 5,200, with an 845 standard deviation. The number of accidents is distributed normally. What is the probability that there will be?

(a) fewer than 6,000 accidents annually.

(b) accidents between 3,700 and 5,000 annually.

Sol: Given mean $\mu = 5,200$ and standard deviation $\sigma = 845$

$$z_1 = \frac{x_1 - \mu}{\sigma} = \frac{6,000 - 5,200}{845} = 0.9467$$

$$P(Z < z_1) = \frac{e^{\sqrt{3}z_1}}{\left(1 + e^{\sqrt{3}z_1}\right)} = 0.8374$$

From Normal distribution method $P(Z < z_1) = 0.8289$.

(c)

$$z_1 = \frac{x_1 - \mu}{\sigma} = \frac{3,700 - 5,200}{845} = -1.7751$$

$$z_2 = \frac{x_2 - \mu}{\sigma} = \frac{5,000 - 5,200}{845} = -0.2367$$

$$P(z_1 < Z < z_2) = \frac{1}{\left(1 + e^{\sqrt{3}z_1}\right)} - \frac{1}{\left(1 + e^{\sqrt{3}z_2}\right)} = 0.3547$$

From Normal distribution method $P(z_1 < Z < z_2) = P(-1.77 < Z < -0.24)$

$$P(z_1 < Z < z_2) = 0.4052 - 0.0384 = 0.3664$$

Problem 3 The preliminary Civil Services Examination (CSE) results in India are distributed normally as of 2023. The standard deviation is 132 and the average score is 625. How likely is it that an individual score will be?

(a) below 700.

(b) Scores from 600 to 750.

Sol: (a) mean $\mu = 625$ and standard deviation $\sigma = 132$ Probability of the score below 700:

$$z_1 = \frac{x_1 - \mu}{\sigma} = \frac{700 - 625}{132} = 0.5682$$

$$\mathrm{P}(Z < z_1) = \frac{e^{\sqrt{3}z_1}}{\left(1 + e^{\sqrt{3}z_1}\right)} = 0.7279$$

From Normal distribution method

$$P(Z < z_1) = P(Z < 0.57) = 0.7279$$

(a) Scores between 600 to 750:

$$z_1 = \frac{x_1 - \mu}{\sigma} = \frac{600 - 625}{132} = -0.1894$$

$$z_2 = \frac{x_2 - \mu}{\sigma} = \frac{750 - 625}{132} = 0.9469$$

$$P(z_1 < Z < z_2) = \frac{1}{\left(1 + e^{\sqrt{3}z_1}\right)} - \frac{1}{\left(1 + e^{\sqrt{3}z_2}\right)} = 0.4188$$

From Normal distribution method $P(z_1 < Z < z_2) = P(-0.19 < Z < 0.95)$.

$$P(z_1 < Z < z_2) = 0.8289 - 0.4247 = 0.4042$$

A hot beverage machine is configured to distribute an average of 250 milliliters each cup. Given a standard deviation of 18 milliliters for drink volume.

(a) how many cups can hold more than 276 milliliters?

(b) What value yields the least 35% of drinks?

Sol:

Mean $\mu = 250$ and standard deviation $\sigma = 18$.

(a) How many of the cups will hold more than 276 milliliters?

$$z_1 = \frac{x_1 - \mu}{\sigma} = \frac{276 - 250}{18} = 1.4444$$

$$P(Z > z_1) = \frac{1}{\left(1 + e^{z_1\sqrt{3}}\right)} = 0.0757$$

From normal distribution table $P(Z > 1.44) = 0.0749$.

(b) At what value do we find the least 35% of drinks?

Given $\mathrm{P}(Z < z_1) = 0.35 \Rightarrow P(Z > z_1) = 0.65$.

$$\Rightarrow \frac{1}{\left(1 + e^{\sqrt{3}z_1}\right)} = 0.65$$

$$\Rightarrow 1 + e^{\sqrt{3}z_1} = \frac{100}{65}$$

$$\Rightarrow z_1 = \frac{ln\left(\frac{7}{13}\right)}{\sqrt{3}} = -0.3574$$

$$z_1 = \frac{x_1 - \mu}{\sigma} = -0.3574 \Rightarrow \frac{x_1 - 250}{18} = -0.3574$$

$$x_1 = 250 - 0.3574 \times 18 = 243.56 \text{ mL}$$

From normal distribution table $P(Z < z_1) = 0.35$.

$$\frac{x_1 - \mu}{\sigma} = -0.38 \Rightarrow \frac{x_1 - 250}{18} = -0.38$$

$$-0.38 \times 18 = x_1 - 250$$

$$x_1 = 250 - 0.38 \times 18 = 243.16 \text{ mL}$$

3. Conclusions

In this essay, I demonstrated that the new probability function follows all of the rules of the general theory of normal distribution. We can easily calculate probability without a table. The definition of probability is chance, hence any differences between the new and original functions should be ignored. Actually, the table values are estimated values of the integral, as the original integral is an improper integral. I strongly feel that this new method has novelty and will be more useful in the future.

Acknowledgements

The author would like to thank the editor and referees for their constructive input and suggestions, which helped enhance this article.

Conflict of interest

The author declares no competing interests.

References

[1] Pitman J. *Probability*. 1st ed. New York: Springer-Verlag; 1993.
[2] Abramowitz M, Stegun IA. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. 10th ed. New York: Dover Publications; 1983.
[3] Lukas E. A characterization of the normal distribution. *Annals of Mathematical Statistics*. 1942; 13: 91-93.
[4] Papoulis A. *Probability, Random Variables and Stochastic Processes*. 4th ed. New York: McGraw-Hill; 1991.
[5] Williams D. *Weighing the Odds: A Course in Probability and Statistics*. Cambridge: Cambridge University Press; 2001.
[6] O'Hagan A. *Kendall's Advanced Theory of Statistics, Vol 2B, Bayesian Inference*. London: Edward Arnold; 2001.
[7] Quine MP. On three characterisations of the normal distribution. *Probability and Mathematical Statistics*. 1993; 14(2): 257-263.
[8] Melnick EL, Tenenbein A. Misspecifications of the normal distribution. *The American Statistician*. 1982; 36(4): 372-373.
[9] Aynes ET. *Probability Theory: The Logic of Science*. Cambridge: Cambridge University Press; 2003.

[10] Wichura MJ. Algorithm AS241: The percentage points of the normal distribution. *Applied Statistics*. 1988; 37(3): 477-484.
[11] Bartlett MS, Kendall DG. The statistical analysis of variance heterogeneity and logarithmic transformation. *Journal of the Royal Statistical Society, Series B*. 1946; 8: 128-138.
[12] Box GEP. Signal to noise ratios, performance criteria and transformations (with discussion). *Technometrics*. 1988; 30: 1-17.
[13] Box GEP, Fung CA. *Studies in Quality Improvement: Minimizing Transmitted Variation by Parameter Design*. Report 8. Madison: University of Wisconsin-Madison, Center for Quality and Productivity Improvement; 1986.
[14] Box GEP, Hunter WG, Hunter JS. *SStatistics for Experimenters*. New York: John Wiley & Sons; 1978.
[15] Brownlee KA. *Statistical Theory and Methodology: In Science and Engineering*. 2nd ed. New York: John Wiley & Sons; 1984.
[16] Carroll RJ, Ruppert D. *Transformation and Weighting in Regression*. New York: Chapman and Hall; 1988.
[17] Norton M, Khokhlov V, Uryasev S. Calculating CVaR and bPOE for common probability distributions with application to portfolio optimization and density estimation. *Annals of Operations Research*. 2019; 299(1-2): 1281-1315.
[18] Das A. A method to integrate and classify normal distributions. *Journal of Vision*. 2021; 21(10): 1.
[19] Lyon A. Why are normal distributions normal? *The British Journal for the Philosophy of Science*. 2014; 65(3): 621-649.
[20] Sun J, Kong M, Pal S. The Modified-Half-Normal distribution: Properties and an efficient sampling scheme. *Communications in Statistics-Theory and Methods*. 2021; 52(5): 1591-1613.